NOTES

Centroids Constructed Graphically

TOM M. APOSTOL
MAMIKON A. MNATSAKANIAN
Project MATHEMATICS!
California Institute of Technology
Pasadena, CA 91125-0001
apostol@caltech.edu
mamikon@caltech.edu

The centroid of a finite set of points Archimedes (287–212 BC), regarded as the greatest mathematician and scientist of ancient times, introduced the concept of center of gravity. He used it in many of his works, including the stability of floating bodies, ship design, and in his discovery that the volume of a sphere is two-thirds that of its circumscribing cylinder. It was also used by Pappus of Alexandria in the 3rd century AD in formulating his famous theorems for calculating volume and surface area of solids of revolution. Today a more general concept, center of mass, plays an important role in Newtonian mechanics. Physicists often treat a large body, such as a planet or sun, as a single point (the center of mass) where all the mass is concentrated. In uniform symmetric bodies it is identified with the center of symmetry.

This note treats the center of mass of a finite number of points, defined as follows. Given n points $\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_n$, regarded as position vectors in Euclidean m-space, relative to some origin O, let $w_1, w_2, \dots, w_n$ be n positive numbers regarded as weights attached to these points. The center of mass is the position vector $\boldsymbol{c}$ defined to be the weighted average given by

$$\boldsymbol{c} = \frac{1}{W_n} \sum_{k=1}^{n} w_k \boldsymbol{r}_k, \tag{1}$$

where W_n is the sum of the weights,

$$W_n = \sum_{k=1}^{n} w_k. \tag{2}$$

When all weights are equal, the center of mass $\boldsymbol{c}$ is called the centroid. If each $w_k = w$, then $W_n = nw$, the common factor w cancels in (1), and we get

$$\boldsymbol{c} = \frac{1}{n} \sum_{k=1}^{n} \boldsymbol{r}_k, \tag{3}$$

which is equivalent to assigning weight 1 to each point. If the points $\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_n$ are specified by their coordinates, the coordinates of $\boldsymbol{c}$ can be obtained by equating components in (1).

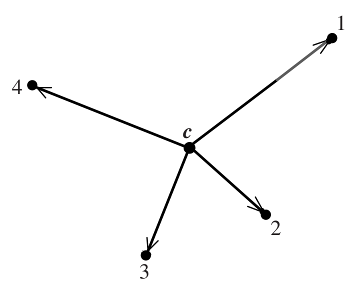
We describe two different methods for locating the centroid of a finite number of given points in 1-space, 2-space, or 3-space by graphical construction, without using coordinates or numerical calculations. The first involves making a guess and forming a closed polygon. The second is an inductive procedure that combines centroids of two disjoint sets to determine the centroid of their union. For 1-space and 2-space both methods can be applied in practice with drawing instruments or with computer graphics programs. Centroids of points in higher-dimensional spaces can be determined with the help of geometric methods by projecting the points onto lower-dimensional spaces, depending on how the points are given. For example, points in 4-space with coordinates (x, y, z, t) can be projected onto points in the xy plane and in the zt plane where graphic methods apply.

Our geometric methods are best illustrated when the weights are equal (computing centroids), and we will indicate in appropriate places how the methods can be modified to the more general case of unequal weights (computing centers of mass).

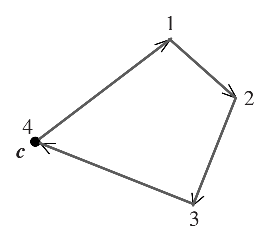
Let $\boldsymbol{c}_k = \boldsymbol{r}_k - \boldsymbol{c}$, the geometric vector from centroid $\boldsymbol{c}$ to $\boldsymbol{r}_k$. Then (3) implies that

$$\sum_{k=1}^{n} \boldsymbol{c}_k = \boldsymbol{O}, \tag{4}$$

so $\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_n$ have centroid $\boldsymbol{O}$. FIGURE 1a shows four points and the geometric vectors represented by arrows emanating from their centroid. FIGURE 1b shows these vectors being added head to tail to form a closed polygon. Of course, the vectors can be added in any order.



(a) Vectors with sum $\boldsymbol{O}$ emanating from the centroid



(b) The vectors form a closed polygon

Figure 1 Vectors from the centroid form a closed polygon when placed successively head to tail

Method 1: Closing a polygon This geometric method for locating the centroid of n given points $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n$ involves choosing a point $\boldsymbol{g}$ as a guess for the centroid. Then we construct a closed polygon and modify the guess once to determine $\boldsymbol{c}$.

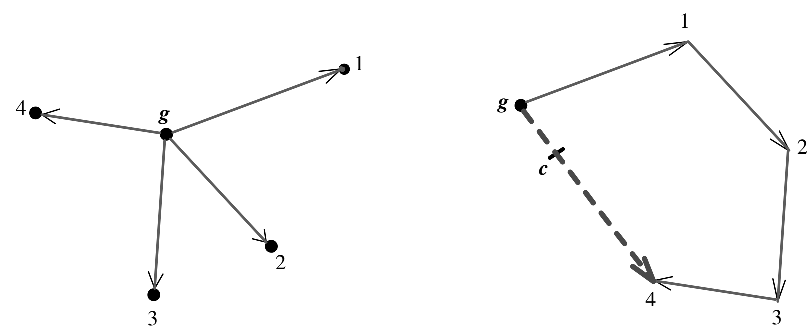
We introduce the deficiency vector defined by

$$\boldsymbol{d} = \sum_{k=1}^{n} (\boldsymbol{r}_k - g) = n\boldsymbol{c} - n\boldsymbol{g},$$

as a measure of the deviation between $\boldsymbol{g}$ and $\boldsymbol{c}$. A knowledge of $\boldsymbol{d}$ gives $\boldsymbol{c}$, because

$$\boldsymbol{c} = \boldsymbol{g} + \frac{1}{n}\boldsymbol{d}. \tag{5}$$

The vector $\boldsymbol{d}/n$ is the error vector. It tells us exactly what should be added to the guess $\boldsymbol{g}$ to obtain the centroid $\boldsymbol{c}$. FIGURE 2 illustrates the method by an example. The four points used in FIGURE 1a are also shown in FIGURE 2a with a guess $\boldsymbol{g}$ for their centroid, and geometric vectors $\boldsymbol{r}_k - \boldsymbol{g}$ drawn from $\boldsymbol{g}$ to each of the four points. For simplicity, in the figure we use labels $k = 1, 2, 3, 4$ to denote these vectors. In FIGURE 2b the vectors are placed successively head to tail, starting from $\boldsymbol{g}$ to form the sum $\boldsymbol{d}$. If a lucky guess placed $\boldsymbol{g}$ at the centroid, the vectors placed head to tail would form a closed polygon as in FIGURE 1b, and $\boldsymbol{d}$ would be zero. But in FIGURE 2a, $\boldsymbol{g}$ is not the centroid, and the polygon formed by these four vectors in FIGURE 2b is not closed. However, an additional vector joining the tail of $\boldsymbol{r}_1 - \boldsymbol{g}$ to the head of $\boldsymbol{r}_4 - \boldsymbol{g}$ will close the polygon. This vector, shown as a broken line in FIGURE 2b, is the deficiency vector $\boldsymbol{d}$. We find $\boldsymbol{c}$ by simply adding the error vector $\boldsymbol{d}/4$ to $\boldsymbol{g}$. In practice, FIGURES 2a and 2b can be drawn on the same graph. We have separated them here for the sake of clarity.



(a) Vectors from guess $\boldsymbol{g}$ to the given points (b) The sum of the vectors in (a)

Figure 2 If $\boldsymbol{g}$ is not the centroid, the polygon obtained by adding the vectors is not closed. It can be closed by adding $-\boldsymbol{d}$ to the other vectors.

Although this example illustrates the method for four points in a plane, the method works equally well for any number of points in 1-space, 2-space, or 3-space. For n points we add the error vector $\boldsymbol{d}/n$ to $\boldsymbol{g}$, as indicated by (5), to get the centroid $\boldsymbol{c}$.

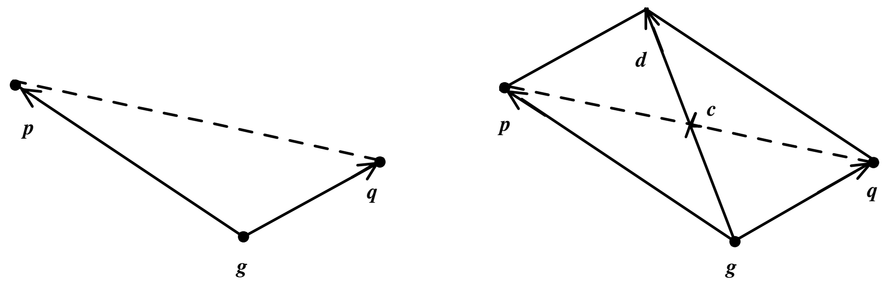
The error vector $\boldsymbol{d}/n$ is easily constructed geometrically. In fact, to multiply $\boldsymbol{d}$ by any positive scalar λ, plot $1/\lambda$ on a number line drawn in a convenient direction not parallel to $\boldsymbol{d}$, and join $1/\lambda$ and the head of $\boldsymbol{d}$ with a line segment. A parallel line to $\boldsymbol{d}$ through the unit on the number line intersects $\boldsymbol{d}$ at $\lambda\boldsymbol{d}$, as is easily verified by similarity of triangles.

Multiplication by a general positive scalar is needed when the method of guessing once is used to find the weighted average $\boldsymbol{c}$ of n given points as defined by (1). As before, we make a guess $\boldsymbol{g}$ and let $\boldsymbol{d} = \sum_{k=1}^{n} w_k(\boldsymbol{r}_k - \boldsymbol{g}) = W_n(\boldsymbol{c} - \boldsymbol{g})$. This gives us $\boldsymbol{c} = \boldsymbol{g} + \boldsymbol{d}/W_n$, and it is easily verified that the foregoing geometric method can be adapted to find $\boldsymbol{d}$, the error vector $\boldsymbol{d}/W_n$, and hence $\boldsymbol{c}$.

Two simple examples illustrate how the method yields familiar interpretations of the centroid.

Example 1: Centroid of two points The centroid of two points is midway between them, and it is instructive to see how the geometric method works in this simple case. FIGURE 3a shows two points $\boldsymbol{p}$ and $\boldsymbol{q}$ and a guess $\boldsymbol{g}$ for their centroid. This may seem to be an outlandish guess because $\boldsymbol{g}$ is not on the line through $\boldsymbol{p}$ and $\boldsymbol{q}$. Nevertheless, the method works no matter where $\boldsymbol{g}$ is chosen. When we add the vector from $\boldsymbol{g}$ to $\boldsymbol{p}$ to that from $\boldsymbol{g}$ to $\boldsymbol{q}$, the sum $\boldsymbol{d} = (\boldsymbol{p} - \boldsymbol{g}) + (\boldsymbol{q} - \boldsymbol{g})$ is one diagonal of a parallelogram, as shown in FIGURE 3b and (5) tells us that centroid $\boldsymbol{c} = \boldsymbol{g} + \boldsymbol{d}/2 = (\boldsymbol{p} + \boldsymbol{q})/2$, which obviously lies midway between $\boldsymbol{p}$ and $\boldsymbol{q}$.

Actually, by choosing $\boldsymbol{g}$ outside the line through $\boldsymbol{p}$ and $\boldsymbol{q}$, there is no need to construct $\boldsymbol{d}/2$, because we know that $\boldsymbol{c}$ lies on this line at the point where $\boldsymbol{d}$ intersects it.



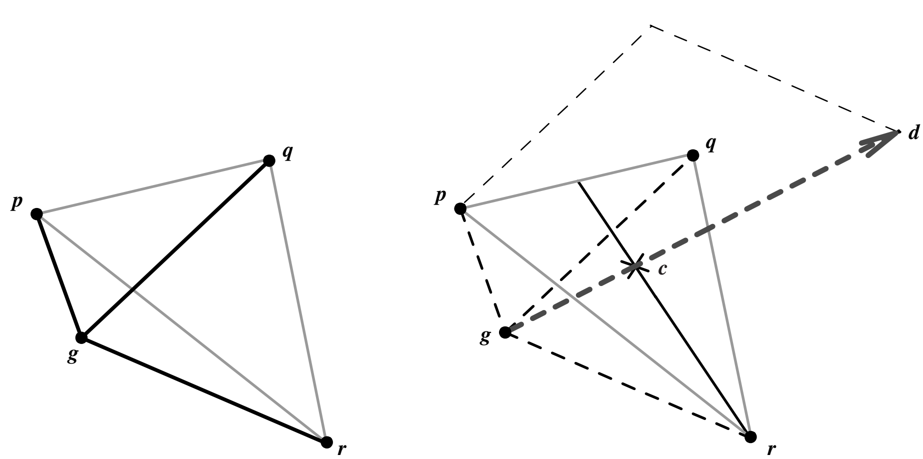
(a) Guess $\boldsymbol{g}$ for the centroid of $\boldsymbol{p}$ and $\boldsymbol{q}$ space (b) Centroid $\boldsymbol{c} = \boldsymbol{g} + \frac{1}{2}\boldsymbol{d} = \frac{1}{2}(\boldsymbol{p} + \boldsymbol{q})$

Figure 3 Determining the centroid of two points

Example 2: Centroid of three points Choose three points $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r}$ and make a guess $\boldsymbol{g}$ for their centroid (FIGURE 4a). The given points are vertices of a triangle, possibly degenerate, but the guess $\boldsymbol{g}$ need not be in the plane of this triangle. According to Method 1, we form the sum

$$\boldsymbol{d} = (\boldsymbol{p} - \boldsymbol{g}) + (\boldsymbol{q} - \boldsymbol{g}) + (\boldsymbol{r} - \boldsymbol{g}) = (\boldsymbol{p} + \boldsymbol{q} + \boldsymbol{r}) - 3\boldsymbol{g}$$

and by (5) the centroid is $\boldsymbol{c} = \boldsymbol{g} + \boldsymbol{d}/3 = (\boldsymbol{p} + \boldsymbol{q} + \boldsymbol{r})/3$. It should be noted that the method works even if $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r}$ are collinear provided we choose $\boldsymbol{g}$ not on the same line, which turns out to be a wise guess in this case. For such a $\boldsymbol{g}$, we know that $\boldsymbol{c}$ is on the line through $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r}$, so vector $\boldsymbol{d}$ will automatically intersect this line at $\boldsymbol{c}$, which means there is no need to divide $\boldsymbol{d}$ by 3.



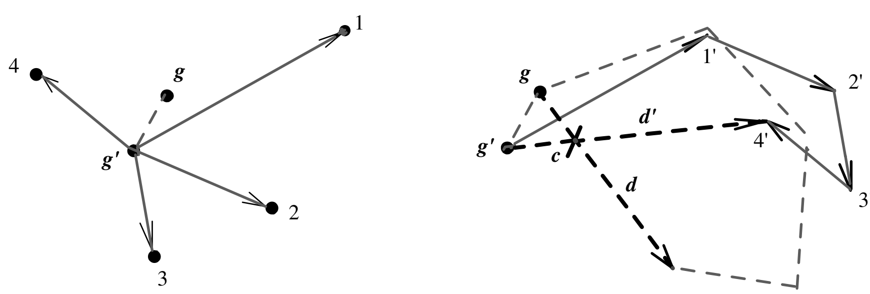
(a) Guess $\boldsymbol{g}$ for the centroid of $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r}$ (b) Centroid $\boldsymbol{c} = \boldsymbol{g} + \frac{1}{3}\boldsymbol{d} = \frac{1}{3}(\boldsymbol{p} + \boldsymbol{q} + \boldsymbol{r})$

Figure 4 The three medians of a triangle intersect at the centroid of the vertices

We can also verify that in the nondegenerate case all three medians of the triangle meet at the centroid, and that the centroid divides each median in the ratio 2 : 1. In

FIGURE 4b the centroid is placed at the origin so that $\boldsymbol{p}+\boldsymbol{q}+\boldsymbol{r}=\boldsymbol{O}$. First consider the median from $\boldsymbol{r}$ to the edge joining vertices $\boldsymbol{p}$ and $\boldsymbol{q}$. The vector from the centroid to the midpoint of this edge is $(\boldsymbol{p}+\boldsymbol{q})/2$ whereas $\boldsymbol{r}=-(\boldsymbol{p}+\boldsymbol{q})$. This shows that the median passes through the centroid and that the distance from the centroid to vertex $\boldsymbol{r}$ is twice that from the centroid to the midpoint of the opposite edge. By interchanging symbols, we find the same is true for the other two medians. So all three medians meet at the centroid, and the centroid divides each median in the ratio 2 : 1 as asserted.

Variation of method 1 that avoids dividing FIGURE 5 shows an alternate way to determine the centroid $\boldsymbol{c}$ of the points in FIGURE 2 that avoids constructing the error vector $\boldsymbol{d}/4$. FIGURE 5a shows the same four points with a new guess $\boldsymbol{g}'$ chosen not on the line through $\boldsymbol{d}$. The polygon of FIGURE 2b appears again in FIGURE 5b as dashed lines, together with a new polygon (solid lines) obtained by adding the vectors $\boldsymbol{r}_k-\boldsymbol{g}'$. This produces a new deficiency vector $\boldsymbol{d}'$ joining the tail of $\boldsymbol{r}_1-\boldsymbol{g}'$ to the head of $\boldsymbol{r}_4-\boldsymbol{g}'$. Because $\boldsymbol{g}'$ is not on the line through $\boldsymbol{d}$, the two geometric vectors $\boldsymbol{d}$ and $\boldsymbol{d}'$ intersect at $\boldsymbol{c}$, as shown by the example in FIGURE 5b.



(a) Vectors from guess $\boldsymbol{g}'$ to the given points (b) Closed polygon formed from guess $\boldsymbol{g}'$

Figure 5 The centroid obtained as the intersection of two deficiency vectors $\boldsymbol{d}$ and $\boldsymbol{d}'$

Although the example in FIGURE 5 treats four points, the method also works for any finite number of points. It should be noted that if the given points are collinear a second guess is not needed provided we choose the first guess $\boldsymbol{g}$ not on the line through the given points. The construction used for three collinear points in Example 2 also works for any number of collinear points. The deficiency vector $\boldsymbol{d}$ will intersect the line through these points at the centroid $\boldsymbol{c}$.

The case of unequal weights is treated similarly as described previously. Make two guesses, and the two deficiency vectors so constructed can be shown to intersect at $\boldsymbol{c}$.

Method 2: Inductive process This method regards the given set of points as the union of two disjoint subsets whose centroids are known or can be easily determined. The centroids of the subsets are combined to determine the centroid of the union. The process depends on how the subsets are selected. For example, we can find the centroid of $n+1$ points if we know the centroid of any n of these points. If $\boldsymbol{c}_n$ denotes the centroid of points $\{1, 2, \ldots, n\}$, so that

$$\boldsymbol{c}_n = \frac{1}{n}\sum_{k=1}^{n}\boldsymbol{r}_k, \quad \text{then} \quad \boldsymbol{c}_{n+1} = \frac{1}{n+1}\sum_{k=1}^{n+1}\boldsymbol{r}_k = \frac{1}{n+1}\left(\sum_{k=1}^{n}\boldsymbol{r}_k + \boldsymbol{r}_{n+1}\right),$$

or

$$\boldsymbol{c}_{n+1} = \frac{1}{n+1}(n\boldsymbol{c}_n + \boldsymbol{r}_{n+1}). \tag{6}$$

In other words, $\boldsymbol{c}_{n+1}$ is a weighted average of the two points $\boldsymbol{c}_n$ and $\boldsymbol{r}_{n+1}$, with weight n attached to $\boldsymbol{c}_n$ and weight 1 attached to $\boldsymbol{r}_{n+1}$. Because $\boldsymbol{c}_{n+1}$ is a convex combination of $\boldsymbol{c}_n$ and $\boldsymbol{r}_{n+1}$ it lies on the line joining $\boldsymbol{c}_n$ and $\boldsymbol{r}_{n+1}$. Moreover, from (6) we find

$$\boldsymbol{c}_{n+1} - \boldsymbol{c}_n = \frac{1}{n+1}(\boldsymbol{r}_{n+1} - \boldsymbol{c}_n),$$

which shows that the distance between $\boldsymbol{c}_{n+1}$ and $\boldsymbol{c}_n$ is $1/(n+1)$ times the distance between $\boldsymbol{r}_{n+1}$ and $\boldsymbol{c}_n$. Repeated use of (6) provides a method for determining the centroid of any finite set.

In the case of unequal weights, (6) becomes

$$\boldsymbol{c}_{n+1} = \frac{1}{W_{n+1}}(W_n\boldsymbol{c}_n + w_{n+1}\boldsymbol{r}_{n+1}),$$

a convex combination of $\boldsymbol{c}_n$ and $\boldsymbol{r}_{n+1}$ that lies on the line joining $\boldsymbol{c}_n$ and $\boldsymbol{r}_{n+1}$. This implies

$$\boldsymbol{c}_{n+1} - \boldsymbol{c}_n = \frac{w_{n+1}}{W_{n+1}}(\boldsymbol{r}_{n+1} - \boldsymbol{c}_n),$$

so the distance between $\boldsymbol{c}_{n+1}$ and $\boldsymbol{c}_n$ is w_{n+1}/W_{n+1} times that between $\boldsymbol{r}_{n+1}$ and $\boldsymbol{c}_n$.

Example 3: Centroid of five points FIGURE 6 shows how this method yields the centroid of five points. Let $\boldsymbol{c}_k$ denote the centroid of the set of points $\{1, 2, \ldots, k\}$. The centroid $\boldsymbol{c}_1$ of point 1 is, of course, the point itself. Using (6) with $n = 1$ we find $\boldsymbol{c}_2$ is midway between 1 and 2. Now connect $\boldsymbol{c}_2$ with point 3 and divide the distance between them by 3 to find the centroid $\boldsymbol{c}_3$. Then connect $\boldsymbol{c}_3$ with point 4 and divide the distance between them by 4 to find $\boldsymbol{c}_4$. Finally, connect $\boldsymbol{c}_4$ with point 5 and divide the distance between them by 5 to get $\boldsymbol{c}_5$. It is clear that this method will work for any number of points.

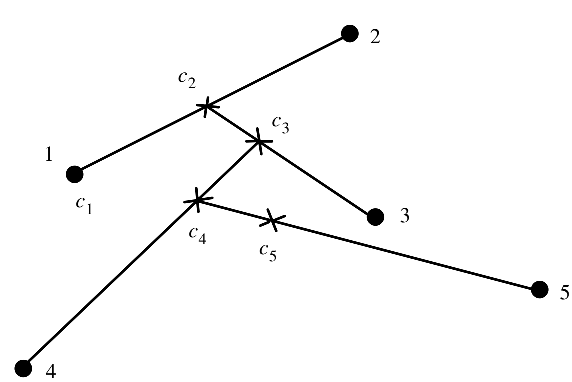


Figure 6 Distance between $\boldsymbol{c}_{k+1}$ and $\boldsymbol{c}_k$ is $1/(k+1)$ times the distance between $\boldsymbol{r}_{k+1}$ and $\boldsymbol{c}_k$

Variation of Method 2 using only bisection and connecting points A variation of Method 2 can be used by those who prefer not to divide vectors into more than two equal parts. This construction uses only two geometric operations—bisection of segments and connecting points—so it applies only to the case of equal weights. We begin with a special example that constructs the centroid using only repeated bisection of segments.

Example 4: Centroid of four points The centroid of four points $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r}, \boldsymbol{s}$, is given by $\boldsymbol{c} = (\boldsymbol{p}+\boldsymbol{q}+\boldsymbol{r}+\boldsymbol{s})/4$. By writing this in the form

$$\boldsymbol{c} = \frac{1}{2}\left(\frac{\boldsymbol{p}+\boldsymbol{q}}{2} + \frac{\boldsymbol{r}+\boldsymbol{s}}{2}\right), \tag{7}$$

we see that the centroid is at the midpoint of the segment joining $(\boldsymbol{p}+\boldsymbol{q})/2$ and $(\boldsymbol{r}+\boldsymbol{s})/2$ which, in turn, are midpoints of the segments from $\boldsymbol{p}$ to $\boldsymbol{q}$ and from $\boldsymbol{r}$ to $\boldsymbol{s}$. By permuting symbols in (7), we see that the centroid is also the midpoint of the segment joining $(\boldsymbol{s}+\boldsymbol{p})/2$ and $(\boldsymbol{q}+\boldsymbol{r})/2$, and of the segment joining $(\boldsymbol{p}+\boldsymbol{r})/2$ and $(\boldsymbol{q}+\boldsymbol{s})/2$. Two quadrilaterals with vertices $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r}, \boldsymbol{s}$ are shown in FIGURE 7, one convex and one not convex. In each case the segment joining $\boldsymbol{p}$ and $\boldsymbol{r}$ (shown dotted) is a diagonal of the quadrilateral with vertices $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r}, \boldsymbol{s}$, as is the segment joining $\boldsymbol{q}$ and $\boldsymbol{s}$. The centroid $\boldsymbol{c}$ lies midway between midpoints of the edges and of the diagonals of the quadrilateral. Any four of the six bisections shown are enough to determine the centroid.

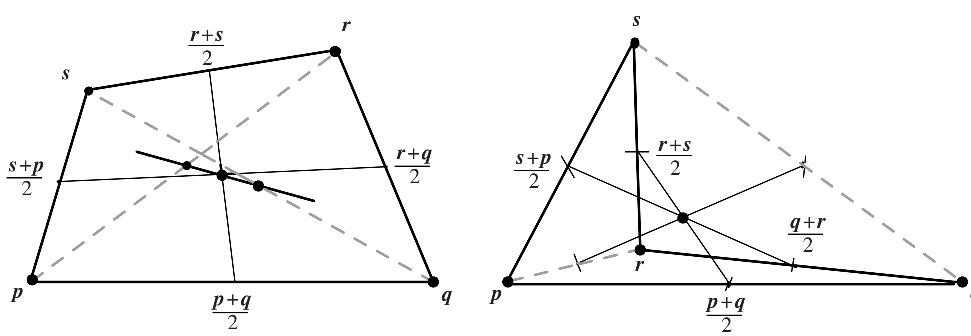


Figure 7 For each set of four points the centroid is midway between midpoints of edges and diagonals

If the four points are collinear, (7) shows that three bisections alone suffice to determine their centroid. An obvious iteration of (7) shows that, if $k \geq 1$, successive bisection suffices to find the centroid of any 2^k points, collinear or not. If the number of points is not a power of 2 we will show that again two operations, bisection of segments and connecting points, suffice to find their centroid. One of the principal tools used in this method is a general property of centroids originally formulated by Archimedes. A general version of this property was used by the authors [**1**] to find centroids of plane laminas. When adapted to finite sets of points, this property can be modified as follows:

ARCHIMEDES'S LEMMA. *If a finite set with centroid* $\boldsymbol{c}$ *is divided into two disjoint sets with centroids* $\boldsymbol{c}_1$ *and* $\boldsymbol{c}_2$*, then the three centroids are collinear. Moreover,* $\boldsymbol{c}$ *lies between* $\boldsymbol{c}_1$ *and* $\boldsymbol{c}_2$.

This is easily proved by the same method we used to obtain (6) in Method 2. Instead of (6) we get a formula of the form

$$c = \frac{1}{n_1 + n_2}(n_1 c_1 + n_2 c_2),$$

where c_1 is the centroid of n_1 points and c_2 is the centroid of n_2, points. This shows that c is a convex combination of c_1 and c_2 and hence lies on the line segment joining them.

It may interest the reader to reconsider Example 2 and see this lemma at work there. The next example shows how Archimedes' Lemma can be used to determine centroids of finite sets of points using only bisection of segments and connecting pairs of points.

Example 5: Centroid of five points FIGURES 8a and 8b show five points distributed in two ways as the union of four points and one point. In each case the centroid of the four points is found as in Example 4, so by Archimedes' Lemma the centroid of all five points lies on the dotted segment joining this centroid with the fifth point. FIGURE 8c shows the intersection of the two dotted segments giving the required centroid of the five points.

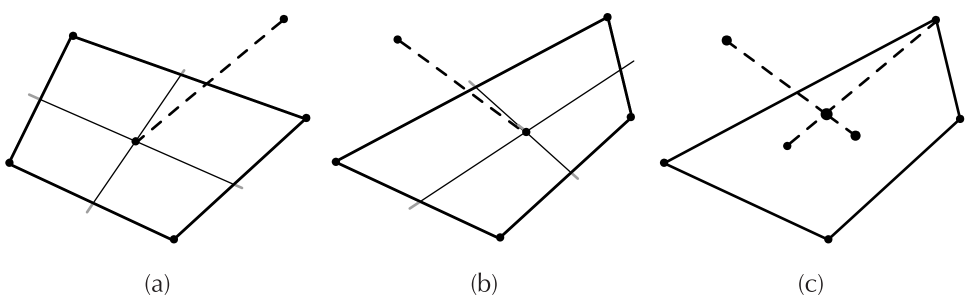


Figure 8 Centroid of five points found by using Archimedes' Lemma twice

In the special case when the five points are collinear, the method cannot give the actual centroid c because the intersecting lines are also collinear. But in this case we can adjoin a sixth point not on the common line and find the centroid c' of the six points by using the Archimedes Lemma twice. Three of the five collinear points together with the sixth point form a quadrilateral whose centroid c_1 can be found as in Example 4. The remaining two of the five collinear points have their centroid c_2 at their midpoint. By Archimedes' Lemma the centroid c' lies on the line joining c_1 and c_2. Now repeat the argument, using a different choice of three of the five collinear points to find another line containing c'. Then the line from the sixth point through c', intersects the line through the five points at their centroid c.

Now we have all the ingredients needed to show that centroids can be determined geometrically using only bisection of segments and connecting pairs of points. We state the result as a theorem, whose proof is constructive and outlines a variation of Method 2.

THEOREM. *For $n \geq 2$, the centroid of n points in m-space can be constructed using only bisection of segments and connecting distinct points.*

Proof. The proof is by induction on n. For $n = 2$ bisection suffices. For $n = 3$ we use bisection and drawing lines in the same manner as described in Example 5. For $n = 4$, bisection alone suffices as described in Example 4. Now suppose the theorem is true for n points, and consider any set of $n + 1$ points. Select one of the $n + 1$ points and join it to the centroid of the remaining n, points which, by the induction hypothesis, has been obtained by bisection of segments and connecting points. Repeat the process, using a different choice for point $n+1$. If all $n+1$ points are not collinear, the two lines so obtained will intersect at their centroid. But if all the $n+1$ points are collinear, choose an additional point outside this line and form, in two ways, a set of n points and a disjoint set of two points (as in Example 5) and apply the inductive procedure twice. This gives two lines whose point of intersection is the centroid, obtained by using only bisection of segments and connecting points. ■

Now we have several procedures at our disposal for finding centroids by an inductive method, two of which have been illustrated for five points. In Example 3 (FIGURE 6) we advanced one point at a time, and in Example 5 (FIGURE 8) we decomposed the set of five points in two different ways as the disjoint union of a single point and a set of four points. In general we can decompose a set of n points into two disjoint subsets in two different ways and use Archimedes' Lemma twice as was done in Example 5. The choice of subsets is a matter of preference, depending on the number of points.

Generalization of a Putnam problem We conclude with an extension of Problem A4 of the 29th William Lowell Putnam Mathematical Competition (1968), which asked to show that the sum of the squares of the $n(n-1)/2$ distances between any n distinct points on the surface of a unit sphere in 3-space is at most n^2. Several solutions of this problem are known, including one by the authors [**2**], using a method that reveals the natural role played by the centroid. The same method is used here to solve a more interesting and more general problem in m-space.

The generalized problem asks for the maximum value of the sum of squares of all distances among n points $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n$ in m-space that lie on concentric spheres of radii $|\boldsymbol{r}_1| \le |\boldsymbol{r}_2| \le \cdots \le |\boldsymbol{r}_n|$. We can imagine a somewhat analogous problem in atomic physics where electrons move on concentric spheres and we ask to minimize the potential energy of the system, which requires minimizing the sum of reciprocals of the distances between charges. Here we wish to maximize the sum of the squares of the distances between points. In [**2**], we showed that this sum is related to their centroid $\boldsymbol{c}$ by the formula

$$\sum_{k<i} |\boldsymbol{r}_i - \boldsymbol{r}_k|^2 = n \sum_{k=1}^{n} |\boldsymbol{r}_k|^2 - n^2 |\boldsymbol{c}|^2, \tag{8}$$

where $\sum_{k<i}$ is an abbreviation for the double sum $\sum_{i=1}^{n} \sum_{k=1}^{i-1}$. Using (8), we can easily maximize the sum on the left, because the right-hand side has its maximum value if and only if $|\boldsymbol{c}|$ reaches its minimum. In other words, locate the points so the centroid is as close as possible to the common center of the spheres. If the value $|\boldsymbol{c}| = 0$ is possible, then $\boldsymbol{c}$ is at the common center (which is chosen also as the origin $\boldsymbol{O}$, and this maximum value is n times the sum of the squares of the radii of the concentric spheres:

$$\max \sum_{k<i} |\boldsymbol{r}_i - \boldsymbol{r}_k|^2 = n \sum_{k=1}^{n} |\boldsymbol{r}_k|^2. \tag{9}$$

For example, in the original Putnam problem, each $|\boldsymbol{r}_k| = 1$, and by locating the points so their centroid is at the center, we find that (9) gives the required maximum:

$$\max \sum_{k<i} |\boldsymbol{r}_i - \boldsymbol{r}_k|^2 = n^2.$$

However, if the points are required to lie on different spheres, the problem of locating them to maximize the sum of squares is more difficult because it is not always possible to place their centroid at the common center. But it can be solved by the graphical method for finding the centroid by closing a polygon. This solution gives a constructive approach as well as a visual interpretation of the results. The solution splits naturally into two cases, depending on how the largest radius $|\boldsymbol{r}_n|$ compares with the sum of all other radii.

Case 1. $|\boldsymbol{r}_n| < \sum_{k=1}^{n-1} |\boldsymbol{r}_k|$, $n \geq 3$. When $n = 3$, this is the triangle inequality, and for $n > 3$ it is a polygonal inequality that makes it possible to choose the vectors $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n$ to have sum zero. In this case we place the origin at the common center of the spheres, and connect the vectors with hinges, head to tail, to form a closed polygon. Because the vectors joining successive edges have sum zero, they can be translated so all initial points are at the origin. The terminal points $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n$ will lie on concentric spheres of radii $|\boldsymbol{r}_1|, |\boldsymbol{r}_2|, \ldots, |\boldsymbol{r}_n|$, their centroid will be at the common center, and the points will satisfy the maximal sum relation (9).

If $n = 3$ the points are vertices of a rigid triangle. But if $n > 3$, there are infinitely many incongruent solutions represented by closed flexible polygons. Any one of these shapes provides a solution: translate each vector parallel to itself to bring all tails to a common point, the common center of the spheres.

Case 2. $|\boldsymbol{r}_n| \geq \sum_{k=1}^{n-1} |\boldsymbol{r}_k|$, $n \geq 2$. In this case the vectors $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n$ cannot have sum zero unless $|\boldsymbol{r}_n| = \sum_{k=1}^{n-1} |\boldsymbol{r}_k|$ and the vectors are on a line. In general, we get the largest possible sum of squares of distances from each other by arranging the vectors along a straight line, with the first $n-1$, vectors $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_{n-1}$ pointing in the same direction, and the nth vector $\boldsymbol{r}_n$ (with the largest radius) pointing in the diametrically opposite direction. Unlike in Case 1, this solution is unique; it gives the largest possible sum of squares of distances from each other consistent with (8), but this largest sum will not reach the maximum provided by the right-hand side of (9) because the centroid is not at the origin.

Where is the centroid? Using the guess $\boldsymbol{g} = \boldsymbol{O}$, we find the deficiency vector $\boldsymbol{d}$ in this case is $\boldsymbol{d} = \sum_{k=1}^{n} \boldsymbol{r}_k = n\boldsymbol{c}$, hence $\boldsymbol{c} = \boldsymbol{d}/n$. Therefore (8) implies that the maximum is given by

$$\max \sum_{k<i} |\boldsymbol{r}_i - \boldsymbol{r}_k|^2 = n \sum_{k=1}^{n} |\boldsymbol{r}_k|^2 - |\boldsymbol{d}|^2, \quad \text{where} \quad |\boldsymbol{d}| = |\boldsymbol{r}_n| - \sum_{k=1}^{n-1} |\boldsymbol{r}_k|,$$

because the vectors are along a line.

As in the original Putnam problem, it is surprising that the maximum in both cases is independent of the dimensionality m of the space if $m \geq 2$. Any solution in one common equatorial plane of the spheres (that is, for $m = 2$) is also a solution in all higher-dimensional spaces.

REFERENCES

1. T. M. Apostol and M. A. Mnatsakanian, Finding centroids the easy way, *Math Horizons*, Sept. 2000, 7–12.
2. ———, Sums of squares of distances in m-space, *Amer. Math. Monthly* **110**, (2003), 516–526.